

# Couplage d'une analyse en composantes principales et d'une approche géostatistique pour l'élaboration de cartes pluviométriques du Centre de l'Algérie du Nord

## Coupling principal component analysis and geostatistics in the creation of rainfall maps for north-central Algeria

Touaïbia Bénina, Inegliz Souhila et Ould Amara Arezki

Volume 19, numéro 3, 2006

URI : <https://id.erudit.org/iderudit/013539ar>  
DOI : <https://doi.org/10.7202/013539ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Université du Québec - INRS-Eau, Terre et Environnement (INRS-ETE)

ISSN

1718-8598 (numérique)

[Découvrir la revue](#)

Citer cet article

Bénina, T., Souhila, I. & Arezki, O. A. (2006). Couplage d'une analyse en composantes principales et d'une approche géostatistique pour l'élaboration de cartes pluviométriques du Centre de l'Algérie du Nord. *Revue des sciences de l'eau / Journal of Water Science*, 19(3), 213–219. <https://doi.org/10.7202/013539ar>

Résumé de l'article

Nous examinons dans cet article, la répartition spatiale de la pluviométrie annuelle en vue de sa cartographie en combinant l'analyse en composantes principales (ACP) et la théorie des variables régionalisées. Deux vecteurs régionaux représentant la tendance pluviométrique la plus probable pour l'homogénéisation des données se sont dégagés. Les pluies annuelles probables de période de retour cinq et dix ans sont évaluées afin de compléter l'information donnée par les pluies annuelles. Les meilleurs modèles de la relation pluie-relief sont aussi recherchés. Par régression multiple, un certain nombre de paramètres morphométriques susceptibles d'expliquer les pluies est identifié, pour leur évaluation en tout point de l'espace. Le variogramme des résidus a présenté une structure spatiale nette. Par krigeage, les paramètres variographiques définis sont calés pour leur interpolation entre les postes pluviométriques. La cartographie automatique a permis d'élaborer les cartes des pluies moyennes annuelles et celles de période de retour cinq et dix ans.

---

# COUPLAGE D'UNE ANALYSE EN COMPOSANTES PRINCIPALES ET D'UNE APPROCHE GÉOSTATISTIQUE POUR L'ÉLABORATION DE CARTES PLUVIOMÉTRIQUES DU CENTRE DE L'ALGÉRIE DU NORD

*Coupling principal component analysis and geostatistics in the creation of rainfall maps for north-central Algeria*

---

*Touaïbia Bénina\*, Inegliz Soubila\*, Ould Amara Arezki\*\**

\*Laboratoire d'Hydrologie, École Nationale Supérieure d'Hydraulique, B.P. 31, 09000 Blida, Algérie

\*\*Service de climatologie, Agence Nationale des Ressources Hydriques, Ravin de la femme sauvage, Alger

Reçu le 6 février 2004, accepté le 9 mai 2005

---

## RÉSUMÉ

Nous examinons dans cet article, la répartition spatiale de la pluviométrie annuelle en vue de sa cartographie en combinant l'analyse en composantes principales (ACP) et la théorie des variables régionalisées. Deux vecteurs régionaux représentant la tendance pluviométrique la plus probable pour l'homogénéisation des données se sont dégagés. Les pluies annuelles probables de période de retour cinq et dix ans sont évaluées afin de compléter l'information donnée par les pluies annuelles. Les meilleurs modèles de la relation pluie-relief sont aussi recherchés. Par régression multiple, un certain nombre de paramètres morphométriques susceptibles d'expliquer les pluies est identifié, pour leur évaluation en tout point de l'espace. Le variogramme des résidus a présenté une structure spatiale nette. Par krigeage, les paramètres variographiques définis sont calés pour leur interpolation entre les postes pluviométriques. La cartographie automatique a permis d'élaborer les cartes des pluies moyennes annuelles et celles de période de retour cinq et dix ans.

**Mots Clés :** *Analyse en Composantes Principales; Géostatistique; Cartographie; Pluviométrie annuelle; Quantiles.*

## ABSTRACT

The spatial distribution of annual rainfall in north-central Algeria was mapped by combining principal component analysis (PCA) and the theory of regionalized variables. The study of the spatial distribution of a natural phenomenon often requires a network of specific measurements. The extension of site-collected and fragmentary data over the whole study area is essential for mapping the variation of the parameter in space. The theory of the regionalized variables is used to account for the structural characteristics of a natural phenomenon in a suitable mathematical form. North-central Algeria, a region covering 390 km from west to east and 360 km from north to south, was the subject of this study. The region contains the Cheliff watershed, coastal Algeria and parts of the Isser and Zahrez watersheds. Ninety-seven rainfall measurement stations were retained, taking into account the quantity and quality of the precipitation data, which covered the time period from 1921/22 to 1960/61 and from 1968/69 to 1998/99, i.e., altogether 70 years but with some gaps. Using PCA, two regional vectors with a multiple correlation coefficient of 0.87, explaining 75% of the total variance, were identified. A linear combination between yearly precipitation

---

\*Auteur pour correspondance :  
Tél./Fax : (213) 25 39 94 47  
Courriel : touaibia@yahoo.fr  
sinegliz@yahoo.fr

and the two regional vectors was established, allowing us to fill the gaps in data among the 97 stations.

Two regional vectors, representing the most likely trend in rainfall used in the homogenization of the data, were identified. A rainfall return period of five to ten years was estimated in order to complete information given by the annual rainfall values. According to the literature, the annual rainfall distribution in Algeria has a positive skewness for which the root normal distribution led to the best fitting of extreme values.

The estimation of the variability in rainfall in the study area required field knowledge. The inadequate distribution of the rainfall network gave a poor representation of the rainfall measurement stations with respect to the altitude. For this reason, a model of the relationship between rainfall and altitude was developed. The altitude was initially analyzed using a digital elevation model (MNT) that samples the elevation using the nodes of a network with a square grid of 2 km x 2 km with the objective of identifying a certain number of morphologic parameters able to explain rainfall variability. The mapped zone was a rectangle oriented from east to west with 35,476 sampled elevations. We chose the model that identifies Lambert co-ordinates (X, Y) and hypsometric (elevation (Z), slope and orientation) parameters. To take into account the orientation of the slopes and the entrenchment of the site, six parameters (TG1, TGE, TGN, TGNE, TGW and TG3) were identified, the calculation of which was based on differences in level. They were identified on the basis of a regular grid (2 km x 2 km) of 25 nodes centered on the rainfall measurement station using a tangent (TG) function. The calculation of smoothed elevation (Zfs) was performed using the TG function and required the construction of a regular grid of nine nodes centered on a rainfall measurement station at an elevation of Z.

The best models of the relationship between rainfall and altitude were identified using multiple linear regressions, by comparing the coefficient of determination ( $r^2$ ) and the sum of residuals. The models obtained were validated by independence and residual normality tests. Some morphologic parameters explaining rainfall variability were also identified, to be used for the estimation of rainfall at any point in the study area using multiple regressions. These models allowed us to extend the series to the nodes (2 km x 2 km) for which the X, Y and Z coordinates are known. A functional relationship between observed and model estimated rainfall was developed. The observed variable «rainfall» can only be accurately defined with the nodes of the grid if the corresponding residuals are established, making the study of the node residuals essential. The knowledge of the regression residuals extended to all grid nodes remains necessary for the mapping of our parameters and allows us to better interpret the resulting variograms. The

interpolation between nodes of the grid (2 km x 2 km) was carried out by kriging, following the investigation of their spacial structure by the calculation of experimental variograms. This spatial structure represented the variance of the differences between the residuals of regression between two distant points of h.

To study anisotropy, the north-south and west-east directions were retained. The exponential model provided a good fit to the regression residuals. The variogram of the residuals has a net spatial structure. Variographic parameters were defined by kriging and fitted for interpolation between rainfall measurement stations. Starting with the elevation grid, as defined by the digital elevation model, the spatial rainfall pattern was obtained by combining the estimated rainfall grids, obtained from the digital elevation model, and the residuals grid. Automatic mapping yielded annual, five-year and ten-year isohyet maps.

**Keywords:** *principal components analysis; geostatistics; mapping; annual rainfall.*

## 1. INTRODUCTION

L'étude d'un phénomène naturel se déployant dans l'espace repose souvent sur un réseau de mesures ponctuelles. Devant une information fragmentaire, recueillie ponctuellement, l'extension des données sur l'ensemble de la région d'étude s'impose pour la représentation cartographique permettant de visualiser ainsi la variation du paramètre dans l'espace. La théorie des variables régionalisées (MATHERON, 1965) est introduite, permettant de rendre compte des caractéristiques structurales du phénomène naturel sous une forme mathématique appropriée. Dans cette optique, une étude sur les précipitations est faite consistant en une analyse des séries pluviométriques, en une étude topographique pouvant influencer les précipitations et en une application à la technique de la cartographie automatique pour leur spatialisation.

## 2. PRÉSENTATION DE LA ZONE D'ÉTUDE

La partie centrale de l'Algérie du Nord a fait l'objet de cette étude. Elle couvre les bassins versants du Cheliff, du côtier algérois, une partie des bassins de l'Isser et de Zahrez. Cette zone offre des particularités climatiques. Elle s'étend sur 390 km d'ouest en est et 360 km du nord au sud (Figure 1).

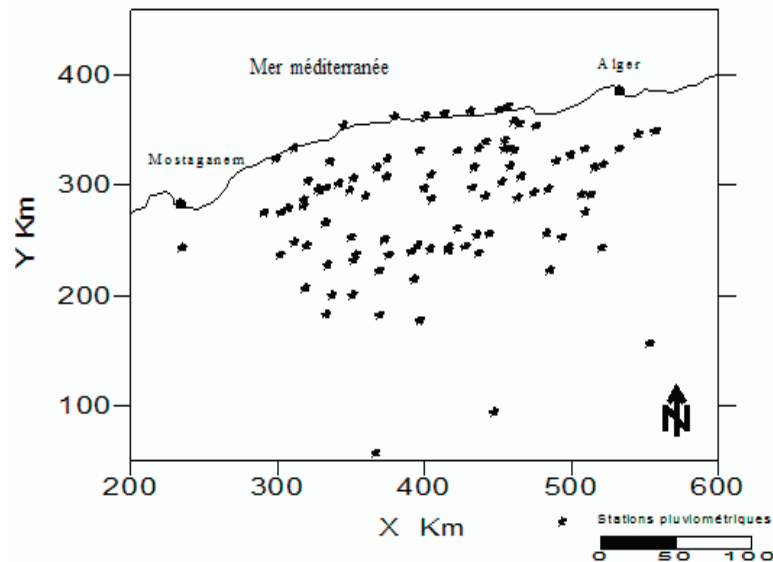


Figure 1 Localisation de la zone d'étude.  
Location of the study area.

### 3. TRAITEMENT DES DONNÉES

Sur les 104 postes pluviométriques répertoriés dans la zone d'étude (ANRH, 2000), 97 ont été retenus sur la base de données fiables en qualité et en quantité, s'étalant sur une période allant de 1921/22 à 1960/61 et de 1968/69 à 1998/99 (70 ans), mais qui reste cependant entachées de lacunes. Les postes observés sont positionnés en figure 1. Les totaux pluviométriques annuels sont analysés.

Deux vecteurs régionaux  $C_1$  et  $C_2$  se sont dégagés de l'analyse en composantes principales (ACP) expliquant plus de 75 % de la variation totale avec un coefficient de corrélation multiple égal à 0,87. Les séries pluviométriques sont ramenées à une combinaison linéaire de ces deux vecteurs par la relation suivante, permettant ainsi le comblement des lacunes pour les 97 stations prises une à une :

$$P_{si} = a_s C_{1i} + b_s C_{2i} + c_s \quad (1)$$

où  $a_s$ ,  $b_s$ ,  $c_s$  sont des coefficients de régression de la station avec les composantes  $C_1$  et  $C_2$  et  $P_{si}$  la pluie manquante de l'année  $i$ .

Sur les 97 séries corrigées et étendues, les pluies moyennes interannuelles, décennales sèches et quinquennales sèches sont évaluées (WOLTING et al., 2000). Il y a lieu de noter que d'après LABORDE et al. (1993), la distribution des pluies annuelles en Algérie présente une dissymétrie positive et la

loi racine normale a donné les meilleurs ajustements pour la détermination des quantiles.

#### 3.1 Construction du modèle pluie - Position géographique

L'estimation de la variabilité des pluies de la région a nécessité la connaissance du champ spatial. Malheureusement, la mauvaise distribution du réseau de mesures pluviométriques liée aux contraintes opérationnelles d'accès au site, notamment en altitude, donne une mauvaise représentativité des postes pluviométriques vis-à-vis du relief, d'où l'élaboration d'un modèle pluie - relief.

Le relief est analysé à partir d'un modèle numérique de terrain (MNT), en échantillonnant les altitudes aux nœuds d'un réseau de mailles carrées de 2 km x 2 km, l'objectif étant de dégager un certain nombre de paramètres morphométriques susceptibles d'expliquer la pluviométrie. Compte tenu des dimensions de la zone à cartographier dans un rectangle Est-Ouest, plus de 35 476 altitudes ont été échantillonnées. Le bloc diagramme de la zone d'étude est présenté à la figure 2.

Notre choix a porté sur le modèle de LABORDE (1982), reposant d'une part sur les coordonnées Lambert (X, Y) qui traduisent la distance, et d'autre part sur les paramètres hypsométriques (altitudes Z et ZfS, pente et exposition T). Pour tenir compte de l'orientation des versants et de l'encaissement du site, six paramètres (TG1, TGE, TGN, TGNE, TGW et TG3) (voir Appendices/Symboles en fin de document) sont

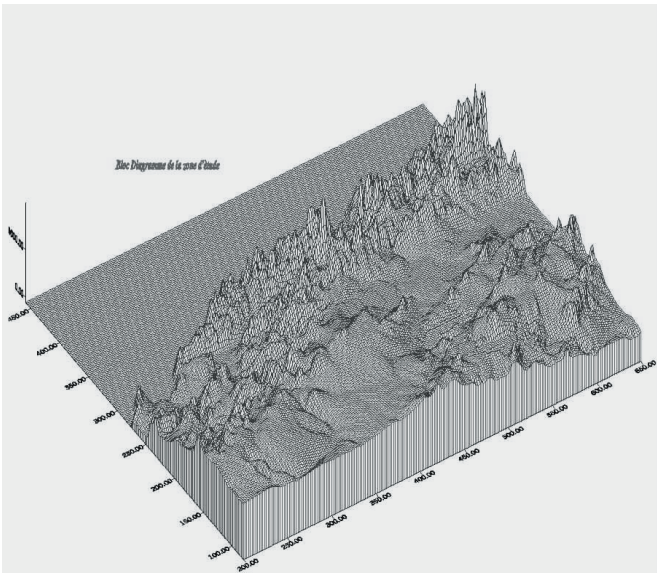


Figure 2 MNT de la zone d'étude.  
Digital elevation model.

identifiés dont le calcul est basé sur la dénivelée. Il s'est fait sur la base d'un maillage régulier de (2 km x 2 km) de 25 nœuds centrés sur le poste pluviométrique en utilisant la fonction Tangente TG (LABORDE, 1986).

$$TG = \operatorname{tg} \theta^{\circ} = \frac{dz}{dl} = \frac{\sum_i^n \alpha_i . \operatorname{tg}(\theta_i)}{\sum_i^n \alpha_i} \quad (2)$$

Le calcul de l'altitude lissée ZfS est donné par la relation (3) et nécessite la construction d'un maillage régulier de neuf nœuds centrés sur le poste pluviométrique d'altitude Z comme le montre la figure 3. Il est recommandé de retenir, pour ce ZfS, la valeur maximale entre l'altitude réelle Z au poste et les quatre autres altitudes obtenues par interpolation linéaire des huit nœuds les plus proches.

$$ZfS = \operatorname{Max} (Z, Z_1 + Z_8/2, Z_2 + Z_7/2, Z_4 + Z_5/2, Z_3 + Z_6/2) \quad (3)$$

Le choix du nombre de paramètres à considérer a été guidé par l'étude des régressions multiples linéaires en comparant, d'une part, le coefficient de détermination R<sup>2</sup> et, d'autre part, la somme des résidus de régression. De toutes les tentatives, la meilleure corrélation des précipitations a porté sur les coordonnées (X, Y, Z) donnant un coefficient de corrélation multiple respectivement égal à 0,74; 0,57 et 0,61 pour la pluie annuelle et les quantiles 10 % et 20 % (DAGNELLIE, 1992). Ces modèles sont validés par les tests d'indépendance et de normalité des résidus R(X,Y,Z) (INEGLIZ, 2002). L'ajustement de ces derniers à la loi racine normale donne de bons résultats pour un seuil de confiance de 5 %. Les modèles retenus sont formulés dans le tableau 1.

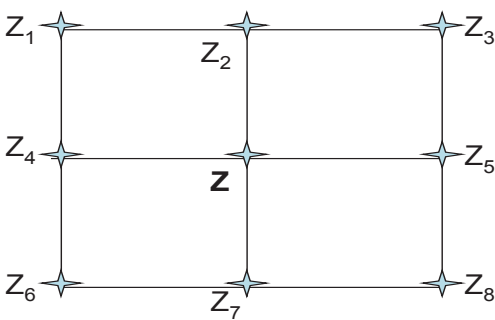


Figure 3 Estimation de l'altitude ZfS.  
Elevation estimation ZfS.

Tableau 1 Relation pluie – relief.  
Table 1 Rainfall – altitude relationship.

Pluie annuelle (mm)	Modèle : P* = f(X,Y,Z) (X et Y en km; Z en m)
Moyenne interannuelle	-0,01 X + 0,71 Y + 0,53 Z + 4,81
Décennale sèche	-0,01 X + 0,71 Y + 0,53 Z + 7,78
Quinquennale sèche	-0,01 X + 0,71 Y + 0,53 Z + 6,42

Les modèles retenus spécifiques (Tableau 1) ont permis d'étendre les séries aux nœuds des mailles (2 km x 2 km), dont on connaît les coordonnées Lambert X, Y et l'altitude Z.

La relation fonctionnelle P(X,Y,Z) = P(X,Y,Z)\* + ε est développée avec P(X,Y,Z), la pluie observée, et P(X,Y,Z)\*, la pluie théorique estimée par le modèle.

La variable « pluie » observée P(X,Y,Z) ne peut être définie fiablement aux nœuds des mailles que si la grille des résidus ε correspondante est établie. De ce fait, l'étude des résidus aux nœuds des mailles s'impose.

3.2 Variographie et interpolation des résidus

La connaissance du résidu de la régression ε étendu à toute la grille reste impérative pour la cartographie de nos paramètres et permet de mieux interpréter les variogrammes obtenus. Leur interpolation aux nœuds de maillage (2 km x 2 km) s'est effectuée par krigeage (OBLED, 1986) après étude de leur structure spatiale par le calcul du variogramme expérimental γ(h). Celui-ci représente la variance des écarts entre les résidus de régression entre deux points distant de h (LABORDE, 1995).



Pour étudier l'anisotropie, deux directions sont retenues, la première Nord-Sud et la deuxième Est-Ouest. Le modèle exponentiel s'adapte bien aux résidus. Il y a lieu de noter que seuls les résidus des précipitations moyennes interannuelles sont pris en exemple dans cet article.

La formulation de  $\gamma(h)$  est donnée par les équations 4 et 5 pour chacune des directions citées et leur représentation graphique respective en figures 4 et 5.

$$\begin{aligned} &\text{Direction Nord-Sud (Azimut } -45^\circ \text{ à } 45^\circ) : \\ &\gamma(h) = 6\,996 (1 - \exp(-h/24)) \end{aligned} \quad (4)$$

$$\begin{aligned} &\text{Direction Est-Ouest (Azimut } -45^\circ \text{ à } 135^\circ) : \\ &\gamma(h) = 6\,983 (1 - \exp(-h/13)) \end{aligned} \quad (5)$$

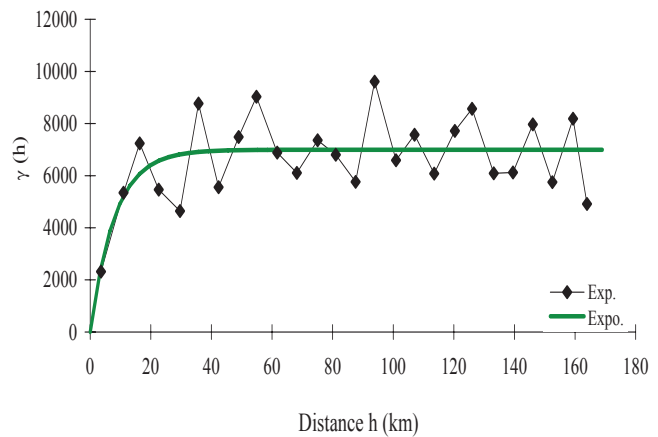
Aussi, les grilles de résidus sont établies pour les deux autres quantiles, la formulation de leur variogramme est donnée dans le tableau 2.

### 3.3 Constitution de la grille finale et de la cartographie automatique

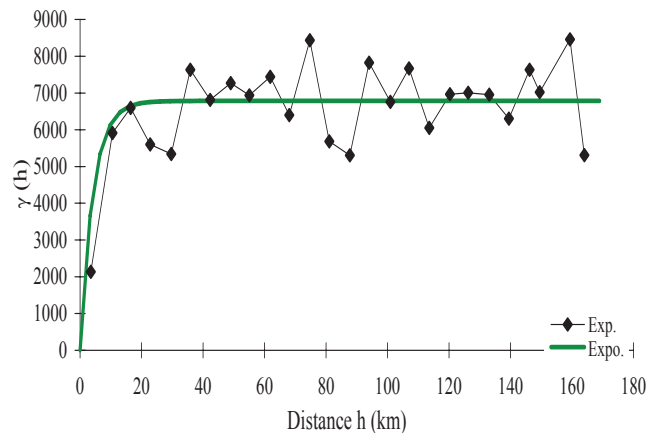
La grille des altitudes étant définie par le modèle numérique de terrain, la grille des pluies dans l'espace est obtenue en combinant les grilles des pluies estimées à partir du modèle numérique de terrain et de la grille des résidus. Cette démarche a été appliquée pour l'élaboration de trois grilles spécifiques, permettant de tracer les courbes isohyètes. Donc, trois cartes pluviométriques sont établies pour un intervalle de confiance de 95 %; il s'agit des cartes des pluies annuelles moyennes, décennale sèche ( $P_{10\%}$ ) et quinquennale sèche ( $P_{20\%}$ ). Les cartes établies sont données en figures 6, 7 et 8.

## 4. CONCLUSION

Le traitement des données, quant à leur homogénéisation et leur spatialisation, est une tâche très délicate. L'approche adoptée pour ce travail reste la meilleure pour notre zone d'étude. La combinaison de l'analyse en composantes



**Figure 4** Variogrammes (expérimental et théorique) des résidus de régression Direction Nord-Sud (Azimut  $-45^\circ$  à  $45^\circ$ ).  
*Experimental and theoretical variograms of the regression residuals North-South direction (Azimuth  $-45^\circ$  à  $45^\circ$ ).*

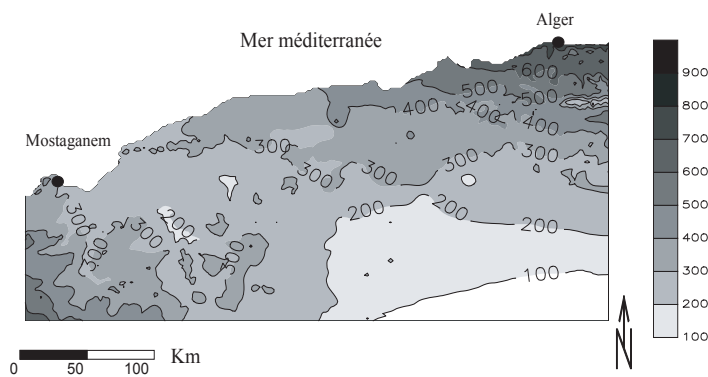


**Figure 5** Variogrammes (expérimental et théorique) des résidus de régression (Direction Est-Ouest ; Azimut  $45^\circ$  à  $135^\circ$ ).  
*Experimental and theoretical variograms of the regression residuals East-West direction (Azimuth  $-45^\circ$  à  $135^\circ$ ).*

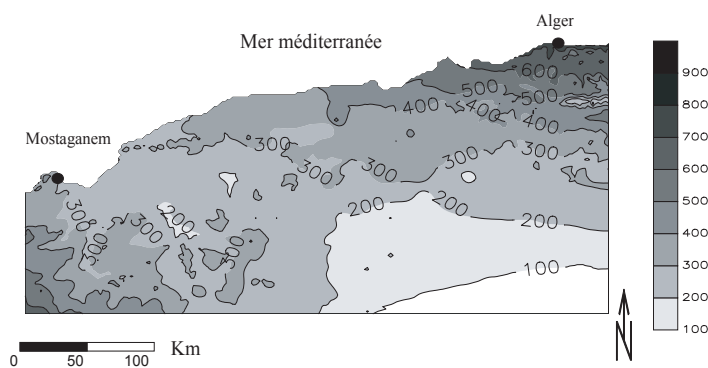
**Tableau 2** Équation des variogrammes  $\gamma(h) = f(h)$ .

**Table 2** Variogram equations  $\gamma(h) = f(h)$ .

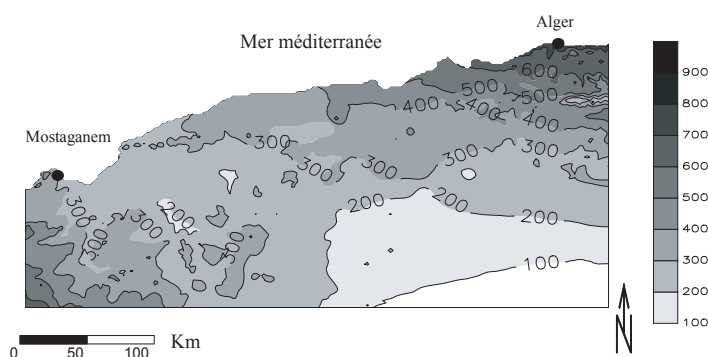
Pluie Direction	Pluie décennale (mm)		Pluie quinquennale (mm)	
	Sèche	Humide	Sèche	Humide
Nord - Sud Azimut ( $-45^\circ$ à $45^\circ$ )	$3\,513 (1 - \exp(-h/15))$	$14\,074 (1 - \exp(-h/28))$	$4\,280 (1 - \exp(-h/20))$	$10\,916 (1 - \exp(-h/27))$
Est - Ouest Azimut( $45^\circ$ à $135^\circ$ )	$3\,495 (1 - \exp(-h/9))$	$14\,000 (1 - \exp(-h/12))$	$4\,288 (1 - \exp(-h/10))$	$10\,737 (1 - \exp(-h/12))$



**Figure 6** Pluies annuelles moyennes (mm).  
*Mean annual rainfall (mm).*



**Figure 7** Pluie annuelle de probabilité de retour 10 ans (Pan, 10 %) (mm).  
*Dry 10-year return period annual rainfall (mm).*



**Figure 8** Pluie annuelle de probabilité de retour 5 ans (Pan, 20 %) (mm).  
*Dry 5-year return period annual rainfall (mm).*

principales pour l'homogénéisation des données et la théorie des variables régionalisées pour la répartition spatiale des pluies ont fourni des résultats très satisfaisants. L'étude des résidus de régression a permis de compenser les biais lors des régressions et d'améliorer ainsi la qualité de l'information. Les cartes d'isohyètes obtenues peuvent être utilisées en absence de données locales ponctuelles. Ce travail, de par son aspect pratique, constitue une approche méthodologique raisonnable

aboutissant à des résultats très satisfaisants. Les cartes élaborées des pluies interannuelles et celles de périodes de retour cinq et dix ans peuvent constituer un point de départ pour une étude sur les pluies de courtes durées dans les zones démunies d'informations pluviographiques.

## RÉFÉRENCES BIBLIOGRAPHIQUES

ANRH, 2000. Données pluviométriques. Document interne. Alger, Algérie.

DAGNELLIE M., 1992. Analyse statistique à plusieurs variables. Vol. 3. Presses agronomiques de Gembloux, Belgique. 362 p.

INEGLIZ S., 2002. Contribution à l'étude de la répartition spatiale de la précipitation dans la région centre de l'Algérie du Nord. Thèse de maîtrise en aménagement hydraulique. Laboratoire d'Hydrologie, École Nationale Supérieure de l'Hydraulique. Blida, Algérie. 120 p.

LABORDE J.-P., 1982. Cartographie automatique des caractéristiques pluviométriques : prise en compte des relations pluviométriques-morphométrie. *Houil. Blanc.*, 4, 330-338.

LABORDE J.-P., 1986. Méthodes d'interpolation et géostatistiques pour la cartographie automatique à l'usage des géographes. Laboratoire d'Analyse Spatiale, Université de Nice-Sophia Antipolis, Nice, France.

LABORDE J.-P., MOUHOUS N. et A. OULD AMARA, 1993. Carte pluviométrique de l'Algérie du Nord à l'échelle 1/ 500 000. Note explicative. Agence Nationale des Ressources en Eau. Projet PNUD/ALG/88/021. Alger. 49 p.

LABORDE J.-P., 1995. Eléments d'hydrologie de surface. Tomes 1, 2 et 3. Cours virtuel. Laboratoire d'Analyse Spatiale, Université de Nice-Sophia Antipolis. Nice, France.

MATHERON G., 1965. Traité de géostatistique appliquée. Tomes 1 et 2. Édition Tecnip, BRGM, Paris, France. 333 p et 171 p.

OBLED C., 1986. Introduction au krigeage à l'usage des hydrologues. Actes des deuxièmes journées hydrologiques de l'ORSTOM, France. (à compléter, date)

WOLTING G., BOUVIER C.H., DANLOUX J. et J.M. FRITZ, 2000. Regionalization of extreme precipitation

distribution using the principal components of the topographical environment. *J Hydrol.*, 233, 86-101.

## APPENDICES/SYMBOLLES

TG : Fonction tangente.

TG1 : Tangente moyenne sous laquelle on voit l'horizon à partir des 25 nœuds les plus proches et entre les azimuts 0° et 360°.

TGW : Tangente moyenne sous laquelle l'horizon Ouest est vu.

TG3 : Tangente moyenne sous laquelle l'horizon Nord-Ouest est vu.

TGN : Tangente moyenne sous laquelle l'horizon Nord est vu.

TGE : Tangente moyenne sous laquelle l'horizon Nord-Est est vu.

TGNE : Tangente moyenne sous laquelle l'horizon Est est vu.

ZfS : Altitude lissée.